

wherein the semantic map represents a determined meaning of the content displayed on the screen, and the screen index indicates a determined position of the content displayed on the screen.

7. The screen navigation apparatus of claim 6, wherein the screen index comprises at least one of the following items: coordinates, grids, and identification symbols, and

the screen analyzer determines at least one of a type, size, and position of the screen index to be displayed on the screen by taking into account at least one of the following factors:

coordinates of the screen index, a screen resolution, and positions and distribution of key contents on the screen, and displays the screen index on the screen based on the determination.

8. The screen navigation apparatus of claim 7, wherein in response to a user selecting one of screen indices displayed on the screen by a user's speech, eye-gaze, or gesture, or any combination thereof, the command composer is configured to interpret the voice command based on screen position information that corresponds to the selected screen index.

9. The screen navigation apparatus of claim 3, wherein the command receiver is configured to receive the input voice command from a user in a predetermined form or in a form of natural language.

10. The screen navigation apparatus of claim 9, wherein the processor further comprises the command receiver.

11. The screen navigation apparatus of claim 9, wherein the command composer comprises a command converter configured to refer to a command set database (DB) and convert the input voice command into a command executable by the screen navigation apparatus.

12. The screen navigation apparatus of claim 11, wherein the command set DB comprises a common command set DB or a user command set DB, or both,

wherein the common command set DB stores common command sets and the user command set DB that stores command sets personalized for a user.

13. The screen navigation apparatus of claim 3, wherein the command composer comprises an additional information determiner configured to determine whether the input voice command is sufficient to be composed into the command, and a dialog agent configured to present a query to request the user to provide additional information in response to the determination indicating that the voice command is not sufficient.

14. The screen navigation apparatus of claim 13, wherein the dialog agent is configured to create the query as multi-stage subqueries, and present a subquery based on a user's reply to a subquery presented in a previous stage.

15. The screen navigation apparatus of claim 3, wherein the command composer is configured to interpret the incoming voice command in stages and compose a command for each stage while the user's voice command is being input, and

the command executer is configured to navigate the screen in stages by executing the commands.

16. The screen navigation apparatus of claim 3, wherein the navigation of the screen comprises one or more of the

following operations: keyword highlighting, zoom-in, opening a link, running an image, playing video, and playing audio.

17. The screen navigation apparatus of claim 1, wherein the screen navigation apparatus is a smartphone, a laptop, a tablet, a smart watch, or a computer, and further comprises a screen and a user interface.

18. A screen navigation method comprising:

receiving a voice command regarding navigation of a screen;

interpreting the voice command based on an analysis result of content displayed on the screen and composing a command; and

performing navigation of the screen based on execution of the command.

19. The screen navigation method of claim 18, further comprising:

analyzing content displayed on the screen and generating a content analysis result.

20. The screen navigation method of claim 19, wherein the content analysis result comprises a semantic map or a screen index, or both,

wherein the semantic map represents a determined meaning of the content displayed on the screen and the screen index indicates a determined position of the content displayed on the screen.

21. The screen navigation method of claim 19, wherein the composing of the command comprises:

in response to the screen index displayed on the screen being selected by a user's speech, eye-gaze or gesture, or any combination thereof, interpreting the received voice command based on screen position information that corresponds to the selected screen index.

22. The screen navigation method of claim 18, wherein the receiving of the voice command comprises receiving the input voice command from a user in a predetermined form or in a form of natural language.

23. The screen navigation method of claim 22, wherein the composing of the command comprises comparing the input voice command to a command set database (DB) and converting the input voice command into the command.

24. The screen navigation method of claim 18, wherein the composing of the command comprises determining whether the input voice command is sufficient to be composed into a command, and in response to a result of the determining being that the voice command is not sufficient, presenting a query to request the user to provide additional information.

25. The screen navigation method of claim 24, wherein the presenting of the query comprises creating the query as multistage subqueries, and presenting a subquery based on a user's reply to a subquery presented in a previous stage.

26. The screen navigation method of claim 18, wherein the composing of the command comprises:

interpreting the incoming voice command in stages while a user's voice command is being input and composing a command for each stage, and the performing of the navigation comprises navigating the screen in stages by executing the commands.

* * * * *